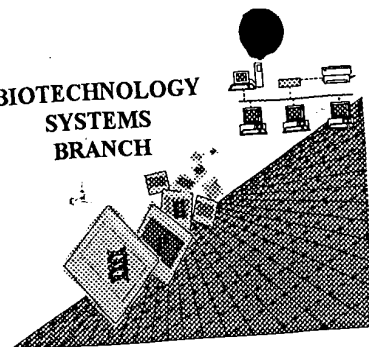


## RAW SEQUENCE LISTING ERROR REPORT

BIOTECHNOLOGY  
SYSTEMS  
BRANCH



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/904,420

Source: OIPK

Date Processed by STIC: 8/8/2001

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER  
VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND  
TRADEMARK OFFICE WEBSITE. SEE BELOW:

### Checker Version 3.0

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 - 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:

<http://www.uspto.gov/web/offices/pac/checker>

# Raw Sequence Listing Error Summary

SERIAL NUMBER: 09/904,420

## ERROR DETECTED

## SUGGESTED CORRECTION

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1      Wrapped Nucleics  
    Wrapped Aminos  
The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2      Invalid Line Length  
The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3      Misaligned Amino  
    Numbering  
The numbering under each 5<sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4      Non-ASCII  
The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5      Variable Length  
Sequence(s)      contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6      PatentIn 2.0  
    "bug"  
A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s)     . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7      Skipped Sequences  
    (OLD RULES)  
Sequence(s)      missing. If intentional, please insert the following lines for each skipped sequence:  
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
This sequence is intentionally skipped  
Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8      Skipped Sequences  
    (NEW RULES)  
Sequence(s)      missing. If intentional, please insert the following lines for each skipped sequence.  
<210> sequence id number  
<400> sequence id number  
000
- 9      Use of n's or Xaa's  
    (NEW RULES)  
Use of n's and/or Xaa's have been detected in the Sequence Listing.  
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10      Invalid <213>  
    Response  
Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11      Use of <220>  
Sequence(s)      missing the <220> "Feature" and associated numeric identifiers and responses.  
Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12      PatentIn 2.0  
    "bug"  
Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13      Misuse of n  
n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

OIPE

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/904,420

DATE: 08/08/2001  
TIME: 07:56:04

Input Set : A:\Rpa1003.app  
Output Set: N:\CRF3\08082001\I904420.raw

3 <110> APPLICANT: Begovich, Ann B  
4 Erlich, Henry A  
5 Grupe, Andrew  
6 Noble, Janelle A  
7 Peltz, Gary A  
8 Reynolds, Rebecca L  
9 Walker, Karen M  
10 Zangenberg, Gabriele  
12 <120> TITLE OF INVENTION: TCF-1 Nucleotide Sequence Variation  
14 <130> FILE REFERENCE: RPA1003  
16 <140> CURRENT APPLICATION NUMBER: US/09/904,420  
17 <141> CURRENT FILING DATE: 2001-07-12  
19 <150> PRIOR APPLICATION NUMBER: US 60/219,812  
20 <151> PRIOR FILING DATE: 2000-07-21  
22 <160> NUMBER OF SEQ ID NOS: 9  
24 <170> SOFTWARE: PatentIn Ver. 2.1

Does Not Comply  
Corrected Diskette Needed

# ERRORED SEQUENCES

26 <210> SEQ ID NO: 1  
27 <211> LENGTH: 2855  
28 <212> TYPE: DNA  
29 <213> ORGANISM: Homo sapiens  
31 <400> SEQUENCE: 1  
32 ggatcccggg ggtcccgggg gccggcgccg gggcccgccg cgaggccgag gtgagccccc 60  
33 gccggcgccg gctcctcccc cgcgggtcgcc gcccgcgccg cccagttgc gcgcccgcct 120  
34 cgggggtctcc agacagagcg tccctgcccc gccgtcgccg ccgacccccg cgggtcccacc 180  
35 gccctcaact cccctccggg tctccctcca ggctctcggg cgggaacacc gtgcgcagag 240  
36 actcttcccg gacaaacttc cagagcccct ggaggacggt gagtttctgc ccggcccggc 300  
37 ttcccttcgt cgcgctcagg ccctggcctc ggtgggacgg ggacgccaaag gaccgcgggg 360  
38 agccgggtgc ctccccacc gcagctcagg aggcggcaga acccaggggt ggagagtggg 420  
E--> 39 ~~gggcccgtt~~ ~~cccggggcgc~~ ~~gccgggtcga~~ ~~gtcacttcgc~~ ~~gtgccctgac~~ ~~ctttatagga~~ 480  
40 gtaaacagac ccccgccatc cccgctccc ctctgcccc ggtgactgac taatccgccc 540  
41 ccttcaggag acagaattgg ccaaggtttc ttggttggag ggtgggggtt gggaggtcaa 600  
42 gtaggggcca cctcggggag gcctgcccct caggctcctc ccctaaaact tggcaactgc 660  
43 gatactccca gccggttcc tcccaagtca ggaacttgca ggggacctct tggcaattct 720  
44 ttttctctca agagcagaca gccttcagtc ccagccgctg ccagggctgg tgtgtctgac 780  
45 ccagctgtgg tttttccagg cctgaaggcc ccggagtga ccagcggcat gtacaaagag 840  
46 accgtctact cgccttcaa tctgctcatg cattaccac cccctcggg agcagggcag 900  
47 cacccccagc cgcagcccc gctggtaagt ggaccccgcc actcaccac cctccttctc 960  
48 atttttcagc acaaggccaa tcagcccccc cacggtgtcc cccaactctc tctctacgaa 1020  
49 catttcaaca gcccatatcc cacccttgca cctgcggaca tcagccagaa gcaaggtaca 1080  
50 agcctgggat gccactcac tcagcttctc tctctgcag ttcacaggcc tctgcagacc 1140  
51 cctgacctct ctggcttcta ctccctgacc tcaggcagca tggggcagct cccccacact 1200  
52 gtgagctggg gagtgtgggc ccagctcagt gttaacttct ttctgcctc caggttcacc 1260  
53 caccatcct tgatgctagg ttctggtgta cctggtcacc cagcagccat cccccaccg 1320

See  
item 9  
on Error  
Summary  
Sheet

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/904,420

DATE: 08/08/2001  
TIME: 07:56:04

Input Set : A:\Rpa1003.app  
Output Set: N:\CRF3\08082001\I904420.raw

```

54 gccattgtgc cccctcagg gaagcaggag ctgcagccct tgcaccgcaa cctgtgagtg 1380
55 aaagacaatc ctgaacaatc tggatttgtg cccctcagga agacacaagc agagtccaag 1440
56 gcagagaagg aggccaagaa gcccaaccatc aagaagcccc tcaatgcctt catgctgtac 1500
57 atgaaggaga tgagagccaa ggtcattgca gagtgcacac ttaaggagag cgctgccatc 1560
58 aaccagatcc tgggcccagc ggtgagacca tgggcagggt ggctggcagg gatgctcccc 1620
59 gaccatcttc agcctggtgc agcctgctga ctccctgatg caccacacct gcccctcttc 1680
60 cctgttgagc tggcacgcgc tgtcgcgaga agagcaggcc aagtactatg agctggcccc 1740
61 caaggagagg cagctgcaca tgcagctata cccaggctgg tcagcgcggg acaactacgt 1800
62 gagtgcctag tgtctgagca tccctccttt tgttccctgc aggggaagaa gaagaggcgg 1860
63 tcgagggaaa agcaccaaga atccaccaca ggtgagacct tctctcgctc taccctctg 1920
64 gcatggctgt gacagacccc tggctcgctt aagaaatgcc gtgctcgctt tggcctcaac 1980
65 cagcagacgg attggtgtgg tccgtgcagg tgggtttgtc cccaggggaa gttctattcc 2040
66 attcattcca tcagagacaa actggcccag agaactcaag gatggtaatg gacaagagtc 2100
67 actgtccatg tcttcttctt ctageccagc ttgaggactg ggatggctgg gcaaggaagc 2160
68 cataggcatt gcggcccctt gccttggtgc agatgtgagt cccacaaaca catctggaga 2220
69 agctcaaagg ccgggactgg gagatgactc ccttggaaga caggagagat gactcccttg 2280
70 gaagacagat gacagcccat aggcctagtg acaaaaggcc cctttgggac cttgtggctg 2340
71 ttctgggaac tgcacctgtc ctaggctctg gccagaccaa gcagaatggc agtctgagga 2400
72 cactgactta ccaccaagt cccaggaaga gaggacaagg aatcagccag gcctgtgcaa 2460
73 aggcagcatt ttttggttgt ggtgtatgac tatgaattca ccctctgtt acagataact 2520
74 ctcttcaact ttctaggag gaaaaagaaa tgcattcggg acttaccggg agaaggccgc 2580
75 tgccccagcc ccgttccttc cgatgacagt gctctaggct gccccgggtc cccagctccc 2640
76 caggactcac cctcatacca tctgctgccc cgcttcccca cagaactgct tactagccct 2700
77 gaaaaagatt attgtagtgt tcaaaatatt tttgtattgt taatgcatca tcatagaaaa 2760
78 acttttaaac atgagaataa agatactttt tactgggttt gtttttcaaa gcctgacctt 2820
79 gaggaataag ctgtttcagt aacagagcat gatata 2855

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/904,420

DATE: 08/08/2001

TIME: 07:56:05

Input Set : A:\Rpa1003.app

Output Set: N:\CRF3\08082001\I904420.raw

L:16 M:270 C: Current Application Number differs, Replaced Application Number  
L:17 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:39 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:1